12-2017

D-Watch: Embracing "bad" multipaths for device-free localization with COTS RFID devices

Ju WANG
Northwest University Xi'an

Jie XIONG
Singapore Management University, jxiong@smu.edu.sg

Hongbo JIANG
Huazhong University of Science and Technology

Xiaojiang CHEN
Northwest University

Dingyi FANG
Northwest University

DOI: https://doi.org/10.1109/TNET.2017.2747583

Published in IEEE/ACM Transactions on Networking,
2017 September,
http://doi.org/10.1109/TNET.2017.2747583

D-Watch: Embracing "Bad" Multipaths for Device-Free Localization With COTS RFID Devices

Ju Wang, *Student Member, ACM*, Jie Xiong, *Member, ACM*, Hongbo Jiang, *Member, IEEE, ACM*, Xiaojiang Chen, *Member, IEEE, ACM*, and Dingyi Fang, *Member, IEEE, ACM*

***Abstract*— Device-free localization, which does not require any device attached to the target, is playing a critical role in many applications, such as intrusion detection, elderly monitoring and so on. This paper introduces D-Watch, a device-free system built on the top of low cost commodity-off-the-shelf RFID hardware. Unlike previous works which consider multipaths detrimental, D-Watch leverages the "bad" multipaths to provide a decimeter-level localization accuracy without offline training. D-Watch harnesses the angle-of-arrival information from the RFID tags' backscatter signals. The key intuition is that whenever a target blocks a signal's propagation path, the signal power experiences a drop which can be accurately detected by the proposed novel P-MUSIC algorithm. The proposed wireless phase calibration scheme does not interrupt the ongoing data communication and thus reduces the deployment burden. We implement and evaluate D-Watch with extensive experiments in three different environments. D-Watch achieves a median accuracy of 16.5 cm for library, 25.5 cm for laboratory, and 31.2 cm for hall environment, outperforming the state-of-the-art systems. In a table area of 2 m$\times$2 m, D-Watch can track a user's fist at a median accuracy of 5.8 cm. D-Watch is also capable of localizing multiple targets which is well known to be challenging in passive localization.**

***Index Terms*— Device-free localization, AoA, multipath.**

I. Introduction

WHILE GPS localization has gained a huge success in outdoor environment, we witnessed an ever-increasing roll-out of indoor location-based applications such as shop navigation [30], [37], augmented reality [19], [61] and room occupancy detection [40] in recent years. Diverse technologies have been explored for localization purposes including infrared [18], [20], camera [27], [29], acoustic [13], [59], visible light [32], [39], Wi-Fi [22], [26], [54], RFID [45], [46], [57], etc. Among all the technologies, radio frequency (RF)-based localization is considered most promising due to its ubiquity and low cost. In particular, RFID is evolving as a major candidate for localization and tracking objects [12], even identifying the material type and imaging the shape of a target in indoor environment [47]. It is no exaggeration to say that we are almost surrounded by RFID tags in our daily lives. These tags are widely used in the bus cards, car keys, clothing security tags, etc. One main reason for this widespread deployment is the simplicity and extremely low cost (5–10 cents USD) of the tags.

The RF-based localization schemes can be further categorized into device-based [22], [51], [53], [54], [58] and device-free [6], [24], [35], [48], [50]. The device-based schemes require the target to be equipped with a device or attached with a tag capable of emitting or reflecting RF signals. However, device-based localization is not applicable in some scenarios. In intruder detection, the targets will deliberately discard any device that can be tracked. In elderly care [57], old people are usually reluctant to hold mobile devices, wear wearables or be attached with RFID tags. These real-life scenarios motive the needs of device-free localization which does not require any device to be attached to the target. On the other hand, device-free localization is usually more challenging as the weaker reflected signals are employed for localization. For example, RFID adopts backscattering strategy for communication and the signal backscattered from the RFID tag is relatively weak. If the backscattered signal gets reflected again from the human target, it becomes extremely weak and it is thus challenging to retrieve this subtle signal for localization. Also multi-target tracking is not a problem for device-based localization but becomes challenging for device-free systems as the reflected signals from multiple targets are all mixed together.

Most RF-based device-free systems [9], [17], [50], [60] employ the received signal strength indicator (RSSI) or channel state information (CSI) for localization which are coarse in accuracy and require labour intensive offline training to obtain the fingerprint database. The fingerprints also need to be updated if there are changes in the environment such as furniture movements, making these systems less realistic for real-life deployment. RSSI/CSI model-based schemes [33], [44], [49], [55] are proposed later to mitigate the fingerprint collection load but suffer from low accuracy in rich-multipath environments. In recent years, motivated by the radar array system [16], [40], [54], AoA-based schemes [7], [48], [57], [61] become popular with the opportunity of multiple antennas attached to a single Wi-Fi access point (AP) or an RFID reader. Wi-Fi APs nowadays are usually equipped with

Manuscript received January 23, 2017; revised July 11, 2017; accepted August 23, 2017; approved by IEEE/ACM Transactions on Networking Editor X. Zhou. This work was supported in part by the National Natural Science Foundation of China under Grant 61572402, Grant 61672428, Grant 61672427, and Grant 61572219 and in part by the National Research Foundation, Prime Minister's Office, Singapore, under its IDM Futures Funding Initiative. *(Corresponding authors: Xiaojiang Chen; Hongbo Jiang.)*

J. Wang, X. Chen, and D. Fang are with the School of Information Science and Technology, Northwest University, Xi'an 710127, China (e-mail: wangju@nwu.edu.cn; xjchen@nwu.edu.cn; dyf@nwu.edu.cn).

J. Xiong is with the School of Information Systems, Singapore Management University, Singapore 188065 (e-mail: jxiong@smu.edu.sg).

H. Jiang is with the School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: hongbojiang2004@gmail.com).

Digital Object Identifier 10.1109/TNET.2017.2747583

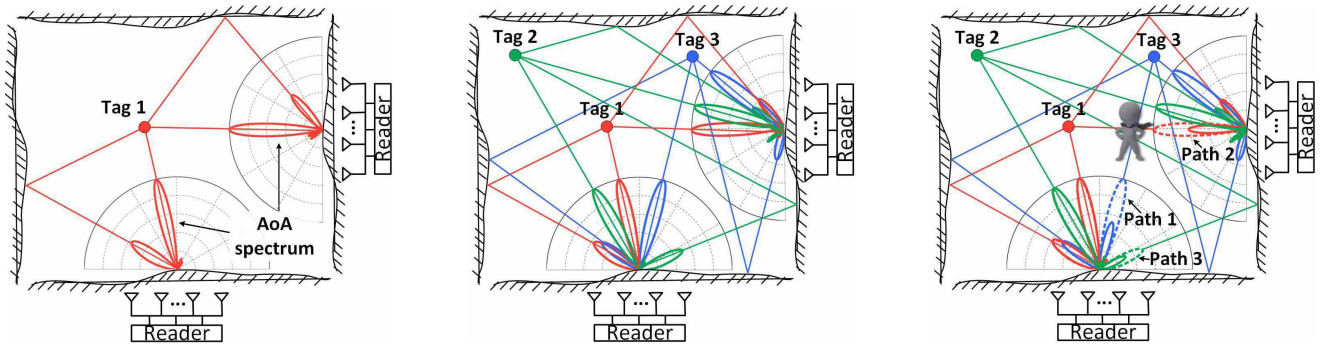


Fig. 1. Illustration of the basic idea of D-Watch. (a) One tag deployed. (b) Three tags deployed. (c) 3 paths are blocked by a target.

an antenna array due to the adoption of MIMO technology [23] in the latest 802.11n and 802.11ac standards. In order to increase the transmission range and accordingly the coverage area, the commodity RFID readers also have multiple antenna ports [4], [46]. The AoA-based schemes are able to achieve a high accuracy without labour intensive offline training. However, one major challenge for the success of AoA-based schemes is the identification of the LoS direct path. This is critical as there are usually rich multipaths indoors and only the direct path signal contains the true angle information of the target. The multipaths are considered detrimental and several works [6], [7], [57] have been proposed to identify or remove these "bad" multipaths.

This paper introduces D-Watch, the first RFID-based device-free localization system that efficiently leverages the "bad" multipath signals to provide decimeter level accuracy. D-Watch is built on the COTS RFID hardware. To locate the target, D-Watch employs the AoA spectrum as shown in Fig. 1(a) and monitors the AoA peak changes to detect the angle information of the target. Specifically, if a target blocks a signal propagation path between the reader and the tag, the corresponding AoA peak will experience a drop. By monitoring the peak amplitude changes on the AoA spectrum, D-Watch can identify the target's angle information without any tag attached to the target, like a "direction watcher (D-Watch)". We then combine the angle information from at least two non-collinear readers to localize the target by the triangulation method.

To further illustrate D-Watch's basic idea, Fig. 1 shows a toy example with two readers and three tags. We can see in Fig. 1(b) that the number of signal paths increases rapidly with two more tags added. When there is a human target in the area, D-Watch localizes the target as illustrated in Fig. 1(c). Specifically, multiple corresponding peaks on the AoA spectra are decreased when the target blocks path 1, path 2 and path 3. By rejecting the wrong angle from path 3 as we will discuss in Section IV-D, D-watch is able to localize the target with the other two paths.

D-Watch efficiently utilizes both the direct path and the reflection paths (multipaths) to identify the angle information of the target. With rich multipaths in indoor environment, D-Watch increases the coverage area significantly, so the deployment density can be well reduced. Further, D-Watch does not need to know the RFID tags' locations since AoA spectrum can be estimated at the reader side without knowing the tags' locations. Thus, the tags can be randomly placed with a high degree of flexibility. D-Watch only requires baseline AoA measurements between tags and readers, which take a few seconds compared with hours measuring and updating the RSSI/CSI signatures at all possible locations in existing fingerprint-based location systems.

Though the idea sounds straightforward, it is non-trivial to realize D-Watch in practice due to the following challenges:

- **Phase calibration:** accurate AoA estimation is highly dependent on the signal's phase value measured at each antenna. However, each radio front end has a random phase offset introduced by the internal oscillator. This random phase offset needs to be carefully addressed before correct AoA information can be obtained. ArrayTrack [52] proposed a wired calibration method by injecting the same signals to the RF front ends with the help of a splitter. However, the wired calibration scheme requires human intervention to plug/unplug the antennas which is time consuming and interrupts the ongoing communication.
- **Signal power estimation of each path:** the signal power of a certain path can not be estimated accurately with the well-known MUSIC algorithm [34] widely used for AoA estimation. The peak amplitude on an AoA spectrum estimated by MUSIC is a probability function [34] and does not have a clear linear relationship with the signal power. When one path is blocked, more than one peak on the MUSIC spectrum may get changed, resulting in a false positive detection. When multiple paths are blocked, MUSIC may only detect one path and miss the other blocked paths.

To deal with the above challenges, we propose a new wireless phase calibration scheme. Compared to the existing methods [36], [52], the proposed scheme does not interrupt the ongoing data communication and runs automatically without the requirement of human intervention. The basic idea is that the estimated AoA angle will match the true angle if the random phase offsets are removed correctly. By deploying tags with known direct path angles and treating the random phase offsets as unknowns, we can compose enough constraint equations to determine the random phase offsets.

We further propose a novel power MUSIC (P-MUSIC) algorithm, which reserves the AoA estimation capability of the traditional MUSIC algorithm and incorporates the signal power estimation capability so both the path angle and path signal power can be obtained at the same time. The key intuition behind P-MUSIC's power estimation is that it applies different weights to the signals received at each antenna so the desired signals at a specific direction add constructively, whereas the signals along other directions add randomly. This alignment boosts the power at the desired direction and averages out the signals along other directions to a small value so the signal power at a specific direction can then be estimated. We build a prototype of D-Watch using four Impinj readers [4] and 21 Alien tags [1]. We evaluate the localization performance in three typical indoor environments: a library, a laboratory and an empty hall corresponding to high, medium and low multipath environment, respectively. D-Watch achieves a median accuracy of 16.5 cm for library, 25.5 cm for laboratory and 31.2 cm for hall, outperforming the state-of-the-art LiFS [44], RASS [60] and RTI [49] systems. For a smaller scale deployment, D-Watch is able to perform fine-grained tracking of a user's fist passively at a median accuracy of 5.8 cm. D-Watch also moves one step further to localize multiple targets simultaneously which is well known to be challenging for passive localization. D-Watch is able to localize three glass bottles at the same time at a maximum error of 17.2 cm as long as they are separated by at least 20 cm from each other.

Contributions: The main contributions of this paper are summarized as follows:

- D-Watch is the first device-free RFID system that efficiently utilizes the "bad" multipaths for localization. D-Watch does not require labour intensive offline training nor needs to know the tags' locations, making D-Watch a promising candidate for real-life deployments.
- The proposed wireless phase calibration scheme outperforms the state-of-the-art wireless calibration method and does not interrupt the ongoing communication. Theoretical and numerical analysis show that we can achieve high AoA estimation accuracies with the calibration method.
- We propose a novel P-MUSIC algorithm, which does not only capture the angle information as the traditional MUSIC algorithm does but also obtains the signal power information of each path.
- D-Watch is implemented on COTS RFID hardware and real-word experimental results demonstrate that D-Watch outperforms the state-of-the-art systems.

II. Background

A. RFID System

An RFID system usually consists of a reader and multiple tags. Tags have no internal battery, so they harvest energy purely from the reader's signal and reply to the reader with a modulated backscatter signal. Typically, a COTS reader is connected to multiple antennas to increase the coverage range. For example, the Impinj xArray reader [4] has 52 antennas and covers a space more than 139 m^2 at a price of 2000 USD.

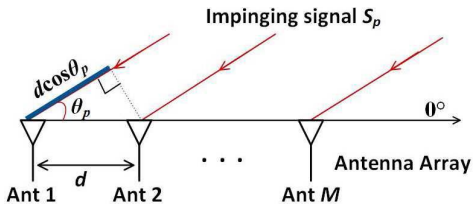


Fig. 2. Phase changes are different when a signal arrives at different antennas of a linear array.

Two points about RFID systems are particularly relevant to localization. First, an RFID tag is extremely cheap at a cost of 5–10 cents USD. RFID reader is relatively expensive. However, most readers have more than two RF ports and are capable of being connected to multiple antennas serving many RFID tags, so the price is well amortized. For example, the ThingMagic Vega Ruggedized RFID Reader at a cost of 875 USD has three antennas. The general purpose Impinj R420 reader at a cost of 1200 USD is capable of supporting four antennas, serving hundreds of tags at the same time. The R420 reader is more expensive than the ThingMagic reader because of the extra HTTP/TCP server function which is not used in our localization system Second, the communication range of RFID today is significantly increased in the last few years to more than 10 meters [4], [45]. Major RFID manufacturers are competing to increase the range [46] and we expect the communication range to continue growing.

B. AoA and MUSIC

AoA Information: When an RF signal propagates in the air, the phase keeps rotating and one wavelength of distance corresponds to a 2π phase rotation. The basic idea of AoA estimation is to measure the signal phase difference at antennas [31] due to different propagation distances. For example in Fig. 2, the adjacent antennas are placed with a space of $d = \lambda/2$ in between, where λ is wavelength. A signal arrives at the antennas along direction θ_p. If the phase measurements at the first two antennas are ϕ_1 and ϕ_2, we can then estimate the AoA θ_p as:

$$\theta_p = \arccos(\frac{\lambda \cdot |\phi_1 - \phi_2|}{2\pi \cdot d}) = \arccos(\frac{|\phi_1 - \phi_2|}{\pi}). \quad (1)$$

AoA Estimation by MUSIC: In reality, Equation (1) does not work because of the multipath signals. MUSIC [34] algorithm is employed for multipath signal AoA estimation. Consider a uniform linear array with M antennas where the first antenna is taken as the reference antenna as shown in Fig. 2. P signals $S = [s_1, s_2, \cdots, s_P]^T$ arrive at the array at directions of $\Theta = [\theta_1, \theta_2, \cdots, \theta_P]$. Since the signals are all mixed, the measured signal x_m at "antenna m" can be expressed as:

$$x_m = \sum_{p=1}^{P} s_p \cdot e^{-j \cdot \omega(m,\theta_p)}, \quad (2)$$

where $\omega(m, \theta_p) = (m-1)\frac{2\pi d}{\lambda} cos(\theta_p)$. The measured signal vector $\mathbf{X} = [x_1, x_2, \cdots, x_M]^T$ at the array is:

$$\mathbf{X} = \mathbf{AS} + \mathbf{n}, \quad (3)$$

where $\mathbf{n}$ is noise, $\mathbf{A} = [\mathbf{a}(\theta_1), \cdots, \mathbf{a}(\theta_p), \cdots, \mathbf{a}(\theta_P)]$ is the steering matrix and $\mathbf{a}(\theta_p)$ is an $M \times 1$ steering vector:

$$\mathbf{a}(\theta_p) = [1, e^{-j\cdot\omega(2,\theta_p)}, \cdots, e^{-j\cdot\omega(M,\theta_p)}]^T, \tag{4}$$

where $(\cdot)^T$ denotes the transpose operation.

MUSIC is based on eigenstructure analysis of the signal vector's correlation matrix $\mathbf{R}$. Based on (3), the correlation matrix $\mathbf{R}$ can be expressed as:

$$\mathbf{R} = \mathbb{E}[\mathbf{X}\mathbf{X}^H] = \mathbf{A}\mathbb{E}[\mathbf{S}\mathbf{S}^H]\mathbf{A}^H + \sigma^2\mathbf{I}, \tag{5}$$

where $(\cdot)^H$ denotes the Hermitian transpose operation, $\mathbb{E}[\mathbf{S}\mathbf{S}^H]$ is the source correlation matrix. The array correlation matrix $\mathbf{R}$ has M eigenvalues $\lambda_1, \cdots, \lambda_M$ associated with M eigenvectors $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_M]$. The largest P eigenvalues correspond to the P incoming signals while the rest $Q = M - P$ correspond to the noise. We follow ArrayTrack [52] to determine the P value based on how many eigenvalues are larger than a fraction (e.g., 2%) of the largest eigenvalue. Based on this process, the corresponding eigenvectors in $\mathbf{U}$ can be classified as signal and noise parts:

$$\mathbf{U} = [\mathbf{U}_S, \mathbf{U}_N] = [\underbrace{\mathbf{u}_1, \cdots, \mathbf{u}_P}_{\mathbf{U}_S}, \underbrace{\mathbf{u}_{P+1}, \cdots, \mathbf{u}_M}_{\mathbf{U}_N}]. \tag{6}$$

We refer to $\mathbf{U}_S$ as the *signal subspace* and $\mathbf{U}_N$ as the *noise subspace*. Due to the orthogonality between the signal steering vector and noise subspace [34], we have:

$$\mathbf{a}(\theta)^H\mathbf{U}_N = \mathbf{0}, \tag{7}$$

when $\theta = \theta_1, \cdots, \theta_P$. Accordingly, the AoA spectrum of MUSIC is given as:

$$B(\theta_p) = \frac{1}{\mathbf{a}^H(\theta_p)\mathbf{U}_N\mathbf{U}_N^H\mathbf{a}(\theta_p)}, \tag{8}$$

which yields sharp peaks at the each signal's AoA.

III. Challenge and Verification

A. Phase Calibration

Accurate AoA estimations are the key part of D-Watch. The AoA estimations are highly dependent on accurate phase measurements. However, a reader's RF front ends introduce random phase offsets into the phase measurements. To examine the amount of phase offsets introduced, we conduct an empirical study over the 16 RF ports on four Impinj R420 readers, with 4 RF ports on each reader. We deploy one tag and one antenna with clear LoS path. The antenna is connected to the 16 RF ports via a same RF cable one by one and the phase measurements are recorded. We select the first RF port as the reference and calculate the phase offsets for the other 15 RF ports. Fig. 3 shows the measured phase offsets ranging from $-85.9°$ to $176°$. These offsets are very random and thus need to be removed in order to achieve reliable AoA estimates. Traditional calibration methods, such as ArrayTrack [52], Argos [36] and Phaser [16], though being able to obtain the phase offsets, require human intervention or take minutes to complete. What is worse, these methods interrupt the ongoing data communication. Note that the phase difference introduced by different tags does not affect the AoA estimation. We use the phase measurements at multiple antennas with respect to *one* tag for the AoA spectrum generation. We never combine phase measurements from multiple tags for the AoA estimates.

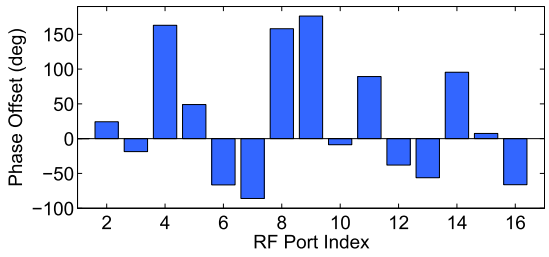


Fig. 3. Random phase offsets at different RF ports.

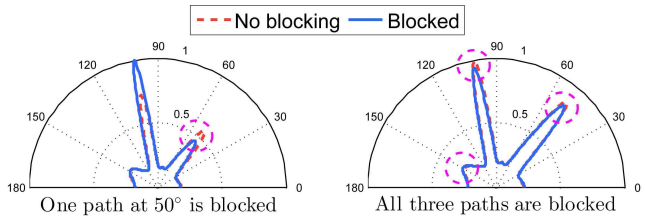


Fig. 4. AoA spectrum change estimated by MUSIC.

B. Limitation of Power Estimation

D-Watch detects the target's direction by observing an obvious signal power reduction when the target blocks the signal path. Thus, an accurate path power change detection is critical for D-Watch's localization. However, the signal power of each path can not be estimated accurately with the well-known MUSIC algorithm [34]. The AoA spectrum estimated by MUSIC is a probability function [34] and the peak amplitude does not represent the signal power level. We run a benchmark experiment to verify this observation. We employ the phase calibration method introduced in Section IV to remove the phase offsets before AoA spectrum estimation. The results are shown in Fig. 4. When one path at 50° is blocked, not just the peak of the blocked path is decreased, the peak amplitudes of other paths may also get changed. That is to say, the power information provided by traditional MUSIC is not reliable for target detection. What is worse, when all three paths are blocked, the amplitudes for all peaks do not change much on the spectrum shown in Fig. 4. Traditional MUSIC will then miss some of the blocked paths and the targets may not even be detected.

IV. System Design

We present the key components of D-Watch in this section followed by a summary of the system workflow.

A. Wireless Phase Calibration

We propose a subspace based wireless calibration met- hod to estimate the phase offsets without interrupting the ongoing communication. The main challenge faced by wireless calibration is the multipath effect. With multiple propagation paths, the signals combine constructively at some antennas while combine destructively at the other antennas. A naive approach

is to remove the multipath signals so that we get the phase measurements only caused by the direct path signal and the RF port phase offsets. We can then obtain the RF port phase offsets because the phase difference introduced by the direct path signal can be calculated. However, in a typical indoor environment, multipaths exist and it's difficult to remove the multipath signals and only keep the direct signal.

In contrast, we propose a novel method which is able to obtain the phase offsets in the presence of multipath signals. The key intuition is the orthogonality property between the signal subspace and noise subspace, if we carry out eigenvalue analysis on the signal correlation matrix [34] as described in Section II-B. Based on this orthogonality property, the product of the noise subspace and signal subspace approaches zero at the LoS angle when the phase offsets are correctly removed. This observation motivates us to find the phase offsets by minimizing the product of noise and signal subspace at the known LoS angle.[1] With phase offsets, we revise the array signal model described by Equation (2) as:

$$\mathbf{X} = \Gamma \mathbf{A} \mathbf{S} + \mathbf{n}, \tag{9}$$

where $\mathbf{X}$, $\mathbf{A}$, $\mathbf{S}$ and $\mathbf{n}$ are the array signal vector, steering matrix, source signal and noise. $\Gamma = \text{diag}\ \{1, e^{j\cdot\Delta\beta_{2,1}}, \cdots, e^{j\cdot\Delta\beta_{M,1}}\}$ is the phase offset diagonal matrix. $\Delta\beta_{m,1} = \beta_m - \beta_1$ refers to the phase offset between "antenna m" and the reference "antenna 1". From Equation (7), we have $\mathbf{a}(\theta)^H \Gamma^H \mathbf{U}_N = \mathbf{0}$ at $\theta = \theta_1, \cdots, \theta_P$. We then estimate the unknown hardware phase offset matrix $\hat{\Gamma}$ by solving the equation $||\mathbf{a}(\theta)^H \Gamma^H \mathbf{U}_N||_{l_2}^2 = 0$ with the angle θ and the noise eigenvector $\mathbf{U}_N$ known. The direct path angle $\theta_{LoS}^{(k)}$ of the k^{th} tag can be easily measured when the locations of the tags and antennas are known.[2] The estimation accuracy of $\hat{\Gamma}$ can be further improved with a larger number of tags. Thus, a more general equation can be given as:

$$\sum_{k=1}^{K} \left\| \mathbf{a}(\theta_{LoS}^{(k)})^H \Gamma^H \mathbf{U}_N^{(k)} \right\|_{l_2}^2 = 0, \tag{10}$$

where K is the total number of tags. Note that the dimensions of $\mathbf{a}(\cdot)$, Γ and $\mathbf{U}_N$ are $1 \times M$, $M \times M$ and $M \times (M-P)$, respectively. By expanding Equation (10), we acquire a number of $K \times (M-P)$ sub-equations. In practice, the number of dominating paths P in the indoor environments is usually smaller than five [52] for each transceiver pair. Thus, we have more than $K(M-5)$ equations for all the tags. While $K(M-5)$ grows in a quadratic fashion, the number of unknown phase offsets $M-1$ grows linearly. This suggests that given enough number of tags, there are enough ($K(M-5) \geq M-1$) equations to determine the unknown $\hat{\Gamma}$.

There are several approaches to solve a set of over-determined equations such as inverting the equations directly or applying the least squares method. However, these approaches are not efficient in solving our problem due to a large amount of non-linear exponential terms in the equations. We formulate an optimization problem to find $\hat{\Gamma}$ that minimize:

$$\hat{\Gamma} = \underset{\Gamma}{\text{argmin}} \sum_{k=1}^{K} \left\| \mathbf{a}(\theta_{LOS}^{(k)})^H \Gamma^H \mathbf{U}_N^{(k)} \right\|_{l_2}^2. \tag{11}$$

The above optimization problem can be solved effectively applying a hybrid method of genetic algorithm (GA) [28] and gradient descent (GD) [8]. Specifically, in each iteration, GA starts initiating all the unknowns and then refines the solution with the GD algorithm to find the closest local minimum.

B. AoA Estimation Accuracy Analysis

In reality, it is not possible to remove all the phase offsets accurately. Here, we analyze the impact of phase calibration error on the AoA estimation accuracy. Suppose there is an error $\Delta\Gamma$ in the phase calibration. Then, the reciprocal of the MUSIC AoA spectrum function is:

$$F(\theta) = ((\Gamma + \Delta\Gamma) a(\theta))^H \hat{U}_N \hat{U}_N^H (\Gamma + \Delta\Gamma) a(\theta), \tag{12}$$

where $\hat{U}_N$ is the noise subspace of the correlation matrix $\mathbb{E}[\hat{\mathbf{X}}\hat{\mathbf{X}}^{\mathbf{H}}]$, which can be obtained by Equation (5) and Equation (6). To simplify Equation (12), we let $\tilde{\Gamma} = \Gamma + \Delta\Gamma$. Due to the phase calibration error, the angle of arrivals $\{\theta_p\}_{p=1}^P$ are distorted by errors $\{\Delta\theta_p\}_{p=1}^P$. Thus, the *p-th* estimated AoA corresponding to the *p-th* signal can be expressed as:

$$\hat{\theta}_p = \theta_p + \Delta\theta_p. \tag{13}$$

By putting (13) into (12), we can get $F(\hat{\theta}_p)$. With the first order Taylor series expansion on $F(\hat{\theta}_p)$, we have:

$$0 = \frac{\partial F\left(\hat{\theta}_p\right)}{\partial \theta} \approx \frac{\partial F(\theta_p)}{\partial \theta} + \frac{\partial^2 F(\theta_p)}{\partial \theta^2} \Delta\theta_p, \tag{14}$$

so, $\Delta\theta_p$ can be described as:

$$\Delta\theta_p \approx -\frac{\frac{\partial F(\theta_p)}{\partial \theta}}{\frac{\partial^2 F(\theta_p)}{\partial \theta^2}}, \tag{15}$$

where,

$$\begin{aligned} \frac{\partial F(\theta_p)}{\partial \theta} = {} & a'(\theta_p)^H \tilde{\Gamma}^H \hat{U}_N \hat{U}_N^H \tilde{\Gamma} a(\theta_p) \\ & + a(\theta_p)^H \tilde{\Gamma}^H \hat{U}_N \hat{U}_N^H \tilde{\Gamma} a'(\theta_p), \end{aligned} \tag{16}$$

and

$$\begin{aligned} \frac{\partial^2 F(\theta_p)}{\partial \theta^2} = {} & a''(\theta_p)^H \tilde{\Gamma}^H \hat{U}_N \hat{U}_N^H \tilde{\Gamma} a(\theta_p) \\ & + 2a'(\theta_p)^H \tilde{\Gamma}^H \hat{U}_N \hat{U}_N^H \tilde{\Gamma} a'(\theta_p) \\ & + a(\theta_p)^H \tilde{\Gamma}^H \hat{U}_N \hat{U}_N^H \tilde{\Gamma} a''(\theta_p), \end{aligned} \tag{17}$$

where a' and a'' represent the first and the second-order derivation of $a(\theta)$ with respect to θ, respectively.

Note that $\hat{U}_N = U_N + \Delta U_N$ and $a(\theta_p)^H \tilde{\Gamma}^H U_N = 0$. Meanwhile, neglecting the derivatives and the perturbation terms of second-order, we get:

$$|\Delta\theta_p| = \frac{\left| \text{Re}\left[-a(\theta_p)^H \tilde{\Gamma}^H \Delta U_N U_N^H \tilde{\Gamma} a''(\theta_p) \right] \right|}{\left\| a'(\theta_p)^H \tilde{\Gamma}^H U_N \right\|^2}, \tag{18}$$

[1] To make sure the LoS path dominates, we place the receiver close to the transmitter with clear LoS path.

[2] The tags' locations are only required for phase calibration. We do not need the tags' locations in localization.

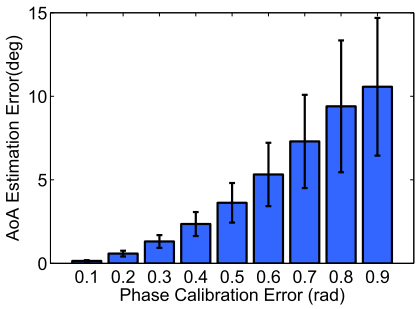


Fig. 5. Impact of phase calibration error on AoA estimation accuracy.

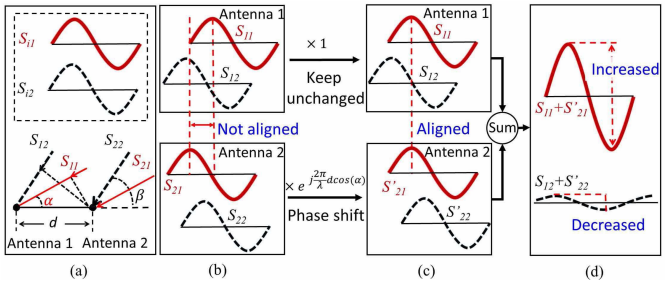


Fig. 6. Illustration of boosting the power of the desired signal in red with alignment and summation. (a) Two signals arrive at two antennas with different AoAs. (b) Two copies of the desired signal in red, namely S_{11} and S_{21}, have a phase shift when they reach at the two antennas. (c) By applying two weights to the measured signals at two antennas, two copies of the desired signal are aligned. (d) The power of the desired signal is boosted when we sum up the two antennas' weighted signals, while the power of another signal will be averagedout to a small value.

According to Khodja and Belouchrani [21], we have:

$$\Delta U_N = -U_S \Lambda_S^{-1} U_S^H \Delta R_{XX}^H U_N, \tag{19}$$

where Λ_S is the signal eigenvalue matrix, and $\Delta R_{XX} = \frac{\left[\hat{X}\hat{X}^H - (\tilde{\Gamma}AS)(\tilde{\Gamma}AS)^H\right]}{G}$, where G is the number of snapshots.

Finally, substituting ΔU_N and ΔR_{XX} into (18), we have:

$$|\Delta\theta_p| = \frac{\left|\mathrm{Re}\left[-a(\theta_p)^H \tilde{\Gamma}^H U_S \Lambda_S^{-1} U_S^H \Delta R_{XX}^H U_N^H \tilde{\Gamma} a'(\theta_p)\right]\right|}{\left\|a'(\theta_p)^H \tilde{\Gamma}^H U_N\right\|^2}. \tag{20}$$

Equation (20) shows the absolute AoA estimation error of the *p-th* signal due to the phase calibration error. To better illustrate the impact of phase calibration error on the AoA estimation accuracy described in Equation (20), we conduct a benchmark simulation with varying amounts of phase calibration errors. Specifically, we generate a phase calibration error for each antenna that follows the Gaussian distribution with a mean value of τ and a variance of 0.1τ, where $\tau = \{0.1 : 0.1 : 0.9\}$. The signals are received at an uniform linear array of eight antennas. Fig. 5 shows that the average AoA error increases with the increase of phase calibration errors. It also shows that the AoA estimation error can be less than $0.15°$ as long as the phase calibration error is no larger than 0.1 radians. In Section VI-A, we show through experiments that our phase calibration method can achieve calibration error less than 0.05 radians. Thus, the proposed phase calibration method does help to achieve high AoA estimation accuracies.

C. Power MUSIC

We introduce a power MUSIC (P-MUSIC) algorithm to address the issue of missing power information in traditional MUSIC algorithm. The key observations are:

- The mixed signal received at one antenna is different from that received at another antenna due to the propagation path difference of signals.
- For a signal at one direction, the phase difference at any two adjacent antennas is the same and can be calculated.
- Different phase weights can be applied to the measured signals at different antennas so the signals at a designed direction add constructively, whereas the signals along other directions randomly add. This alignment boosts the power at the desired direction and averages out the signals at other directions to a small value, so the signal power at the designed direction can be obtained.

Fig. 6 (a) shows a toy example with two antennas spaced at a distance of d. Two signals arrive at two different angles α and β. We donate the two signals as S_{i1} and S_{i2} when they reach "antenna i". The two signals are mixed together at each antenna. To obtain the signal power at angle α, we apply two carefully chosen weights $[1, e^{j\frac{2\pi}{\lambda}d\cos(\alpha)}]$ to the measured signals and then sum them up. S_{11} and S'_{21} then add up constructively, while the other signals add up destructively, as shown in Fig. 6 (d). With the alignment and summation, we successfully reduce the power in other angles and obtain the power at angle α. Note that with just two antennas, the other signals may still have a chance of adding up constructively. However, with more antennas, other signals add up with random phase shifts, which will average out to a small value.

Without loss of generality, assume we want to identify the signal power s_1 along direction θ_1. The measured signal at "antenna m" is x_m. Then the sum of the weighted versions of x_m at direction θ_1 is given as:

$$\begin{aligned}
&\sum_{m=1}^{M} x_m \cdot e^{j\cdot\omega(m,\theta_1)} \\
&= \left(s_1 + \sum_{p=2}^{P} s_p\right) + \left(s_1 + \sum_{p=2}^{P} s_p \cdot e^{-j\cdot[\omega(2,\theta_p)-\omega(2,\theta_1)]}\right) \\
&\quad + \cdots + \left(s_1 + \sum_{p=2}^{P} s_p \cdot e^{-j\cdot[\omega(M,\theta_p)-\omega(M,\theta_1)]}\right) \\
&= M\cdot s_1 + \sum_{p=2}^{P}\left[s_p\left(\sum_{m=1}^{M} e^{-j\frac{2\pi(m-1)d}{\lambda}[\cos(\theta_p)-\cos(\theta_1)]}\right)\right]
\end{aligned} \tag{21}$$

As shown in the above equation, the signal s_1 adds constructively and the amplitude gets increased roughly M times while other signals average out when adding up with random phase shifts of $\frac{2\pi\ (m-1)d}{\lambda}\cos(\theta_p)$. For any interested signal s_p, with the alignment and summation, the power of the designed signal along direction θ_p will be much higher than

the power along other directions with a relatively large M. The signal power along the direction θ_p with the received signals at the M antennas is then given as:

$$P_B(\theta_p) = \|s_p\|^2 \approx \frac{\left\|\sum_{m=1}^{M} x_m \cdot e^{j\cdot\omega(m,\theta_p)}\right\|^2}{M^2}. \quad (22)$$

We then integrate this power information into the traditional MUSIC to acquire both AoA and power estimations. Intuitively, we can simply dot-multiply the power estimation $P_B(\theta_p)$ and the MUSIC AoA spectrum $B(\theta_p)$. However, the MUSIC peak amplitude is a probability value, which distorts the estimated signal power. We solve this problem by designing a normalization function $Nor(B(\theta_p))$ to normalize all the peak amplitudes to "1". We thus remove the peak amplitudes from MUSIC and only keep the angle information of the peaks. Our P-MUSIC function is then given as below:

$$\begin{aligned}\Omega(\theta_p) &= P_B(\theta_p) \cdot Nor(B(\theta_p)) \\ &= \frac{\left\|\sum_{m=1}^{M} x_m \cdot e^{j\cdot\omega(m,\theta_p)}\right\|^2}{M^2 \cdot Nor(\mathbf{a}^H(\theta_p)\mathbf{U}_N\mathbf{U}_N^H\mathbf{a}(\theta_p))},\end{aligned} \quad (23)$$

where M is the number of antennas, $\mathbf{a}(\theta_p)$ and $\mathbf{U}_N$ are the steering vector and the noise subspace eigenvector defined in Section II-B. Note that P-MUSIC does not need to know the value of θ_p. By searching θ_p from 0 to π like the traditional MUSIC, P-MUSIC is able to estimate the signal power along each direction of the signal path.

D. Target Localization

D-Watch combines the identified target angles from several readers to determine the target's location. Suppose ξ readers identify a set of AoA spectra changes $\Delta\Omega(\theta_1), \cdots, \Delta\Omega\ (\theta_\xi)$. To localize a target, we compute the likelihood function $L(O)$ of the target be located at a position O, and take the location estimate with a maximum likelihood as the target location. To compute the likelihood, the basic idea is that a larger AoA peak change indicates a higher probability that the target is located at this angle. The likelihood $L(O)$ is given as:

$$L(O) = \prod_{i=1}^{\xi} \Delta\Omega(\theta_i). \quad (24)$$

We then divide the monitoring area into grids[3] and search for the grid with the highest $L(O)$. The hill climbing scheme is employed to find the most likely target location estimate.

Note that we may retrieve wrong angle information if the target blocks a reflection path before the signal reaches the reflector. In Fig. 1(b), Path 3 is blocked but the angle information detected is not correct. In reality, this wrong angle information can be identified if there is only one target. Because a target cannot block two paths at the same reader at the same time. Whenever we detect multiple blocked paths at one reader, we know only one of the detected angles is pointing to the true location of the target. We further discover that the locations estimated from the wrong angles are distributed at random positions and even far outside of the monitoring area. While, the correct angles will localize the target to close-by positions. We can then apply outlier rejection to identify the wrong angle.

[3]Smaller grid size leads to more accurate results but takes more time to search. We balance this tradeoff and set the grid size as 5 cm×5 cm for the three indoor environments and 2 cm×2 cm for the smaller table area.

E. Putting Things Together

Now we show the workflow of D-Watch.

Step 1 (Data Collection): D-Watch collects a set of baseline AoA data between the tags and the reader when no target is present. Note that this process is very different from the traditional fingerprint database collection which takes hours. This process for D-Watch is just several transmissions between the readers and tags which can be well completed within seconds. D-Watch then acquires another set of measurements when the target moves into the monitoring area.

Step 2 (Data Pre-Processing): D-Watch employs the proposed wireless phase calibration method to remove the phase offsets that exist in the collected data. Note that the calibration process is a one-time effort for one power on/off cycle so D-Watch does not need to carry out phase calibration frequently.

Step 3 (Target Angle Estimation): With proposed P-MUSIC algorithm, D-Watch generates two sets of AoA spectra based on the online data and the baseline data. By comparing the amplitude changes of the AoA peaks, D-Watch can accurately identify the target's angle information at each reader.

Step 4 (Target Localization): By combining the identified angle information from multiple readers, D-Watch is able to obtain the target's location estimate with triangulation scheme.

V. Implementation

Experimental Environments

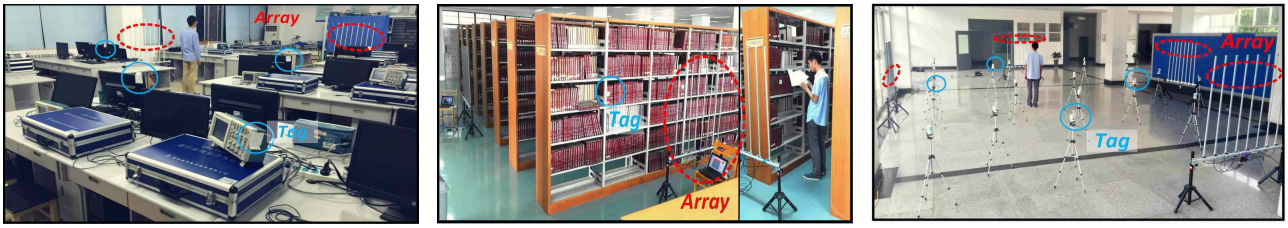
we conduct experiments in three typical indoor environments: a library, a laboratory and an empty hall corresponding to high, medium and low multipath environment. The laboratory with a size of 9 m×12 m has many small objects such as test chambers, displays, etc., as shown in Fig. 7(a). Part of the library area with a size of 7 m×10 m has many book shelves full of books, as shown in Fig. 7(b). The shelf has a height of 2.5 m and is made of metal and wood, resulting in rich multipath and strong NLoS. Part of the hall with a size of 7.2 m×10.4 m is shown in Fig. 7(c).

Human Target and Object Target

We ask the students to act as the human target. To demonstrate the high accuracy of D-Watch in a small scale deployment, we employ three glass bottles full of water as our object targets. The bottle has a bottom diameter of 7.8 cm and a height of 22 cm.

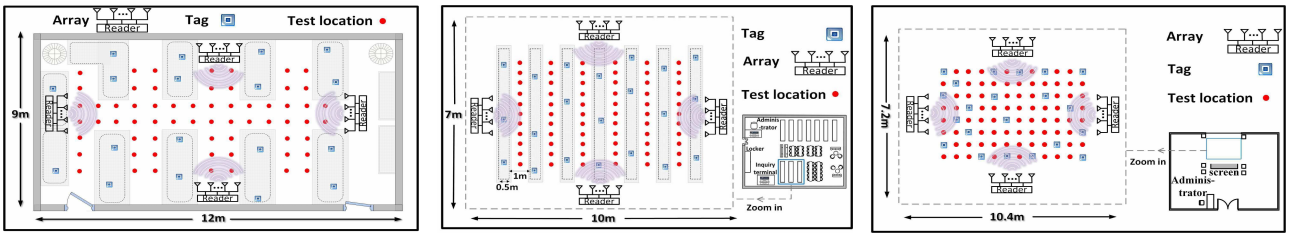
Implementation

(i) *Readers and tags*. We employ four Impinj R420 readers [4] for our experiments without any hardware or firmware modification. Each reader has four RF ports and is equipped with one Impinj GPIO Adapter [4]. An Impinj antenna hub is connected to one of the reader's four RF ports as shown in Fig. 9. Usually, we need 8 antennas for highly accurate AoA



(a) (b) (c)

Fig. 7. Three typical indoor environments corresponding to medium, high and low multipath environment. (a) Laboratory environment. (b) Library environment. (c) Hall environment.



(a) (b) (c)

Fig. 8. Deployment layouts with the positions of arrays, tags and test locations marked. (a) Laboratory deployment layout. (b) Library deployment layout. (c) Hall deployment layout.

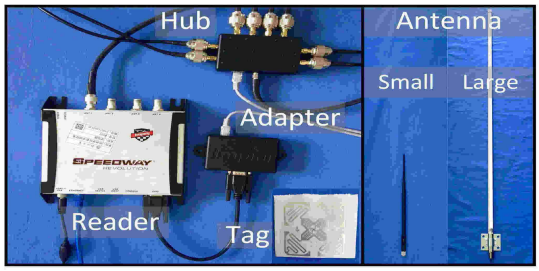


Fig. 9. RFID hardware and antennas employed.

estimation [52]. However, the Impinj Speedway R420 reader only has four antenna ports, thus we use the antenna hub provided by Impinj to support more antennas. The readers are compatible with EPC Gen2 standard [3] and the antennas work in a time division multiplexing mode. The time slot for each antenna is about 200 us [4]. The reader operates in 920.5–924.5 MHz. We use 21 cheap Alien tags [32] as shown in Fig. 9, which cost 1.5 USD in total.

(ii) *Antenna and array*. We employ two different types of antennas for our experiments as shown in Fig. 9. The small one is ANS-900 omni-directional antenna [2] and the large one is Q900F-900 omni-directional antenna [5]. Since RFID devices communicate by backscattering the signal which is usually weak, the small antenna [2] provides a relatively small communication range of 3 m. The large antenna provides a communication range of 12 m. Each linear array is consisted of 8 antennas with a half wavelength space of 16.25 cm between adjacent antennas. The antennas and the RF front ends have been calibrated using the wireless phase calibration method introduced in Section IV-A.

(iii) *Server and algorithm implementation*: The proposed schemes and algorithms are implemented in C# and Matlab. The server is a desktop with 3.6 GHz CPU (Intel i7-4790) and 8 GB memory. The server communicates with the RFID readers using low level reader protocol (LLRP) [14]. All the tags' backscatter packets received at the readers are forwarded to the server through Ethernet cables. The size of a RFID backscatter packet is small since the packet only contains the tag's ID which is 12 bytes at most [4]. The reader does not need to emit signals all the time and a 0.1 s transmission interval is good enough for our localization and does not increase the transmission overheads.

Default Deployment Setup

The deployment layouts of the three environments are shown in Fig. 8. In each environment, we deploy four[4] readers and 21 tags. The locations of the readers are known while the tags are randomly placed without a need to know their locations. We choose 63, 66 and 75 test locations in the laboratory, library and hall, respectively. The test locations are uniformly distributed with a 0.5 m distance in between. The objects attached with tags are usually placed on the table or held in the hand so their heights are between 1 to 1.5 m above the ground. We place the antenna array at a height of 1.25 m. When we evaluate the impact of number of tags and tag-array height difference, we employ more tags and change the default setup. Unless specifically mentioned, we use the default setup for performance evaluation.

Experimental Methodology

When a target moves into the monitoring area, the readers receive 10 backscatter packets from each tag and forward them to the server. At the server side, D-Watch identifies the angle

[4]Note that the number of readers can be reduced to one if we employ coaxial cables to connect four antenna arrays to one reader.

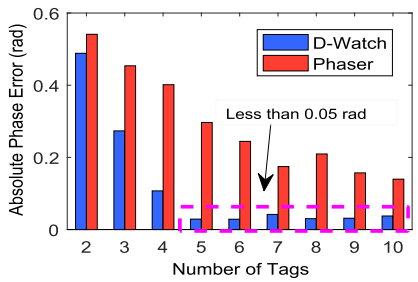


Fig. 10. The proposed phase calibration method is much more accurate.

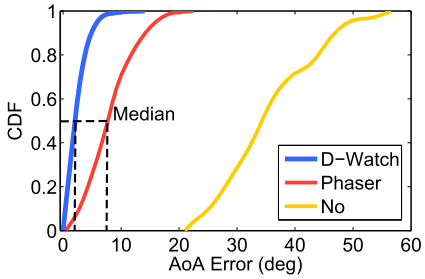


Fig. 11. LoS AoA estimation with proposed calibration method.

information of the target at each reader and combines the information from several readers to obtain the target's location estimate. We repeat the experiments 40 times at each test location.

VI. Performance Evaluation

A. Microbenchmark

We start with two benchmark experiments to validate the effectiveness of the proposed wireless calibration method and P-MUSIC algorithm.

1) Verification of Phase Calibration Method: We measure the random phase offsets caused by the reader's radio front ends and compare our method with the state-of-the-art wireless calibration method proposed in Phaser [16]. We take the phase offsets obtained from the wired calibration method proposed in ArrayTrack [52] as the ground truth. In the laboratory environment, tags are randomly attached to objects located 1–8 m away from the array. We vary the number of tags to estimate the phase offset matrix $\hat{\Gamma}$. The phase offset estimation errors are shown in Fig. 10. The proposed method is able to achieve phase error less than 0.05 radians when more than four tags are employed. It suggests that the proposed method achieves a high calibration accuracy even with a relatively small number of tags. Note that a phase calibration error of 0.05 radians will cause an even smaller AoA error since there are multiple antennas and the calibration errors have positive and negative values which may cancel out each other. We then compare the direct path AoA estimation errors when employing different calibration methods. The results in Fig. 11 show that our calibration method achieves a high AoA estimation accuracy, i.e., a median error of 2°, outperforming the Phaser system.

2) Verification of Power MUSIC: To minimize the influence of multipath, we conduct this experiment in the empty hall environment. We place two laptops with metal shells as reflectors to obtain two controlled reflection signals.

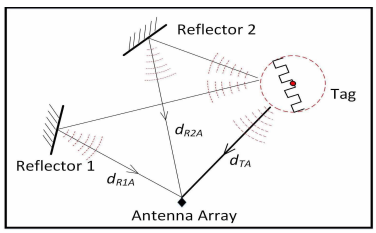


Fig. 12. Deployment layout for Power MUSIC verification.

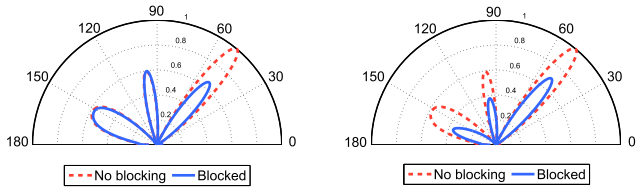


Fig. 13. AoA spectrum changes estimated by P-MUSIC. (a) One path is blocked. (b) Three paths are blocked.

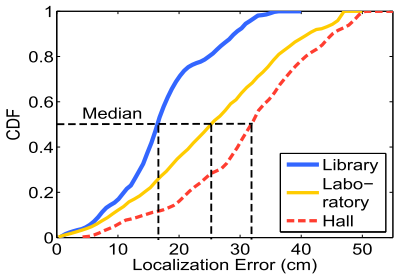


Fig. 14. CDF plot localization error.

The deployment layout is shown in Fig. 12. Three human targets walk around to block the direct path and reflection paths. Fig. 13 illustrates the changes of AoA spectra when we apply P-MUSIC. Compared with traditional MUSIC shown in Fig. 4, the changes of the AoA peaks estimated by P-MUSIC match our expectations, i.e., the blocked path peak experiences a clear drop and the unblocked peaks remain unchanged.

B. Overall Localization Performance

In reality, human target is too big to be treated as a point. As the human targets have a width of 32 cm to 40 cm, we consider there is no localization error as long as the estimation is within a 36 cm range. Otherwise, we calculate the error as the minimum difference between the estimated location and this 36 cm range.

We show the localization performance for the human target in three different environments in Fig. 14. The results in Fig. 14 show that D-Watch achieves the best performance in the library environment with a median and 90% error as small as 16.5 cm and 28.9 cm, respectively. D-Watch's median accuracy slightly decreases in laboratory and hall environments to 25.3 cm and 32.1 cm. In contrast, D-Watch performs even better in the richer-multipath library environment, implying that "bad" multipaths are efficiently utilized to improve the localization performance.

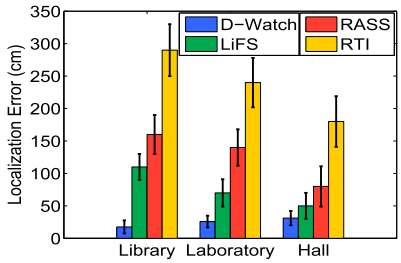


Fig. 15. Localization errors compared by 4 schemes in different indoor environment.

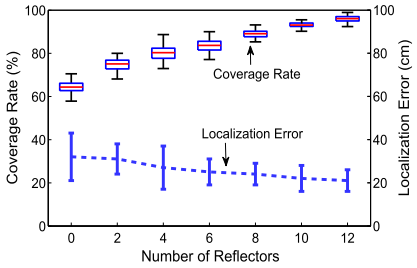


Fig. 16. Localization errors with varying number of reflectors.

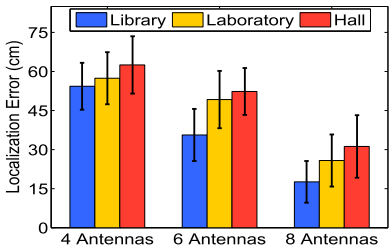


Fig. 17. Localization errors with varying number of antennas.

C. Localization Comparison

We compare D-Watch with three state-of-the-art systems:

- **LiFS:** LiFS [44] is a model-based device-free localization system, which does not require any training. LiFS first models the target location and CSI amplitude measurements of all wireless links as a set of equations, then it localizes the target by solving the equations.
- **RASS:** RASS [60] is a training-based device-free localization/tracking system, which utilizes the RSSI reading as location feature. RASS collects a set of fingerprints, i.e., the RSSI data when a target is located at a location, then it utilizes the support vector regression algorithm to localize the target. In this experiment, we employ the "LIBSVM" tool [25] used in RASS to localize a target.
- **RTI:** RTI [49] is a radio-tomography-based device-free localization system, which also does not require any training. RTI maps the sum of RSSI variances of all wireless links into a two-dimensional space and takes the position with the largest RSSI variance as the target's location estimation. RTI requires the prior knowledge of all the tags' locations. For a fair comparison, we employ the scheme proposed in the well known Tagoram [56] system to localize the tags' positions.

Fig. 15 depicts the localization errors obtained for all four schemes in three different environments. In library environment, D-Watch achieves the best performance with a mean error as small as 17.6 cm. LiFS, RASS and RTI have larger mean errors of 1.1 m, 1.6 m and 2.9 m, respectively. In laboratory environment, D-Watch, LiFS, RASS and RTI have mean errors of 25.5 cm, 0.7 m, 1.4 m and 2.4 m. In Hall environment, D-Watch, LiFS, RASS and RTI achieve mean location errors of 31.2 cm, 0.5 m, 0.8 m, 1.8m, respectively. Fig. 15 demonstrates that D-Watch performs the best in all the environments due to the wireless phase calibration scheme and the power music algorithm proposed.

In the library and laboratory environments, LiFS and RASS perform worse than D-Watch since the CSI amplitude (or RSSI) measurements vary over time and are subject to multipath errors. On the other hand, D-Watch can make use of multipath to improve the location accuracy. The performance of RTI is not good because RTI's localization accuracy is dependent on all tags' precise location information. However, it is difficult to obtain accurate tag location estimation in NLoS environment, resulting in degradation of localization accuracy.

Fig. 15 also shows that the location accuracies of LiFS, RASS and RTI are decreased with more mutipaths. Surprisingly, D-Watch has an improving localization accuracy with more mutipaths, which demonstrates D-Watch can leverage multipath to improve location accuracy as claimed.

D. Impact of Number of Multipaths

Besides the direct path, D-Watch utilizes the multipaths to localize a target. With more multipaths, D-Watch achieves a higher coverage rate[5] and improves the localization accuracy. To demonstrate this, we place up to 12 reflectors such as laptops and metal reflectors to create more multipaths in the hall environment. Fig. 16 shows that the coverage rate gets increased significantly since more propagation paths exist in the monitoring area. It also shows a mean error decrease from 31.2 cm to 20.8 cm as more paths are now restricting the target's location estimate.

E. Impact of Number of Antennas

With more antennas at each reader, D-Watch is able to achieve a finer resolution in AoA estimation and capture more paths, which accordingly increases the localization accuracy as shown in Figure 17. In the library environment, the mean localization error is 54.3 cm for four antennas, 35.6 cm for six antennas and 17.6 cm for eight antennas. The lab and hall environments show similar results.

F. Impact of Number of Tags

With more tags, higher number of signals will be reflected creating more paths to cover the monitoring area and constraint the target's location. In the library environment, we vary the number of tags from 7 to 47 with a step size of 5. The experimental results match our expectations as shown in Fig. 18. So both tags and reflectors can actually increase the multipaths in the environment. Thus, in an indoor environment with more reflectors, the density of the tags can be reduced.

[5]Coverage rate is defined as the number of locations can be localized divided by total number of test locations.

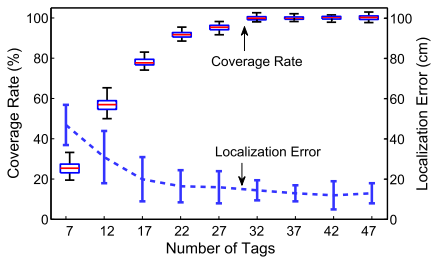


Fig. 18. Localization errors with varying number of tags.

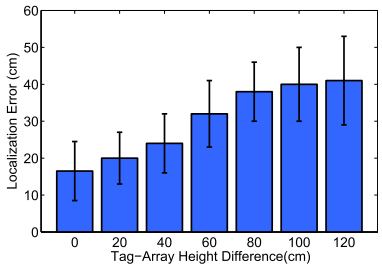


Fig. 19. Localization errors with varying tag-array height differences.

G. Impact of Tag-Array Height Difference

In reality, tags attached on books or laptops are placed on the table or held in the hand with a height of 1–1.5 m above the ground. If the tags and the arrays are not at the same height, we would like to study whether such a height difference will cause significant localization errors. Fig. 19 shows that D-Watch can still achieve a mean localization error of 40 cm even when the height difference is as large as 120 cm. When the height difference is 40 cm, the mean error is 24 cm which is only slightly higher than no height difference.

H. Multi-Target Localization

The high spatial resolution of D-Watch's AoA spectrum enables a fine-grained multi-target localization. The intuition is that a target is not able to block all the paths simultaneously. When multiple targets are located sparsely, each target will block a disjoint subset of paths and thus can be separated and individually located. However, when many targets exist or two targets are too close to each other, it is still challenging to accurately localize each of them.

We study the performance of D-Watch for multi-target localization in this section. To evaluate the granularity of multi-target localization, we employ three glass bottles placed on a 2 m×2 m table as shown in Fig. 21. The bottles are filled with water. We place two small-antenna arrays at the midpoint of the bottom and right of the table. 26 tags are placed at the other two sides. Fig. 20 (a)∼(c) shows three snapshot localization results when the three targets are separated roughly by 130 cm, 50 cm and 20 cm respectively. For each snapshot, we collect 30 data samples and the localization results are mapped into the heatmap, where the red dot represents the location estimate and the black cross indicates the ground truth. D-Watch localizes the three targets accurately with a maximum error of 17.2 cm when they are located sparsely, such as 130 cm and 50 cm apart. When the three targets are close to each other within 20 cm, the three targets have a tendency to merge into one on the heatmap and D-Watch can not localize each individual accurately. We believe with more number of antennas such as 16 and 64, finer AoA resolution can be achieved and we can further increase the accuracy and separability for multi-target localization.

I. Application Example: Tracking Fist in the Air for Virtual Screen Touch

One popular application in the research community is drawing/writing in the air [38] or tracking the human's fine-grained gestures [11]. We briefly investigate the capability of D-Watch on these applications. We let a human user write the characters "P" and "O" along a pre-marked trajectory using his fist in the 2 m×2 m table area as shown in Fig. 22. The user moves his hand at a natural writing speed, i.e., about 0.5 m/s. Fig. 22 shows that the trajectory of the user's fist measured by D-Watch matches the ground truth quite well. To understand the tracking accuracy of D-Watch, we mark a set of continuous test points on the table. Fig. 23 shows that the median tracking error can be as small as 5.8 cm and 9.7 cm when 26 tags and 13 tags are employed respectively.

VII. Discussion

Deadzone Problem

When a target does not block any path, it is in a "deadone" where the target can not be detected. In this case, D-Watch can utilize the mobility of a target to mitigate this problem. A human target moves continuously in space and the target can still be localized before and after entering into the "deadzone". These location information can then be utilized to estimate the target's current location when target is in the "deadzone". Moreover, there are rich multipaths in a typical indoor environment and the tags are very cheap so we can increase the number of tags to reduce the amount of deadzones. Thus, the probability of this extreme case is quite low in reality.

Mobility

Since the human target moves continuously, we can track the target by snapshots. In the indoor environment, the walking speed of a human is around 1–2 m/s. Note that the transmission interval of D-Watch is 0.1 s so the target moves only 10–20 cm in this short period which does not affect D-Watch's performance much. Also Doppler shift can be applied to estimate the target's walking speed to further improve the location accuracy.

Latency of D-Watch

The system latency includes the time to collect packets and the time to calculate the target location. We run our localization algorithm many times and the average processing time is below 100 ms. For each tag, 2–3 packets are collected for localization in D-Watch. The time taken to collect data packets depends on how frequently the packets are sent out. If the packet is sent out every 100 ms, the end-to-end system latency of our system is still below 0.5 s.

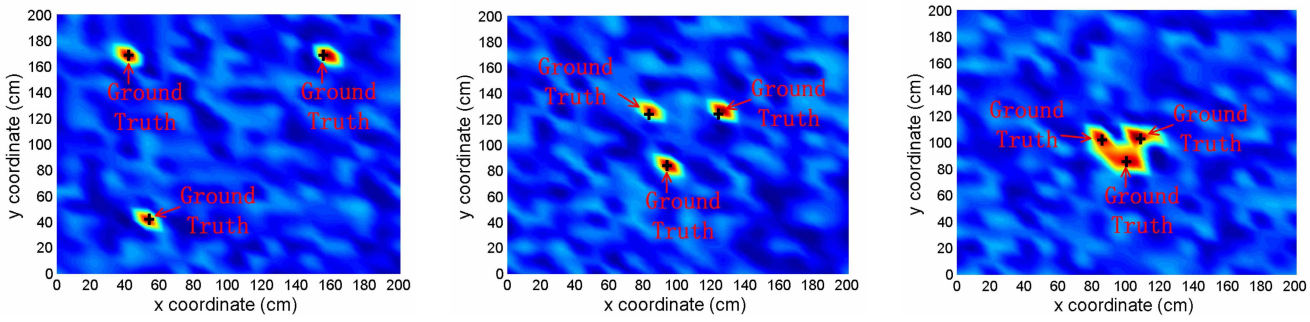


Fig. 20. Performance of multi-target localization. (a)-(c) show the heatmaps of three snapshot localization results. (a) Targets are 130cm apart. (b) Targets are 50cm apart. (c) Targets are 20cm apart.

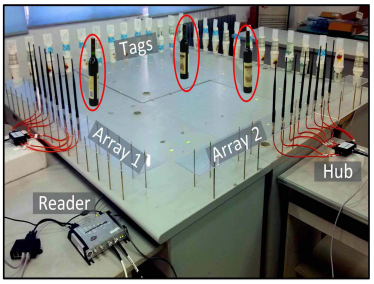


Fig. 21. A 2 m × 2 m table area with three glass bottle targets.

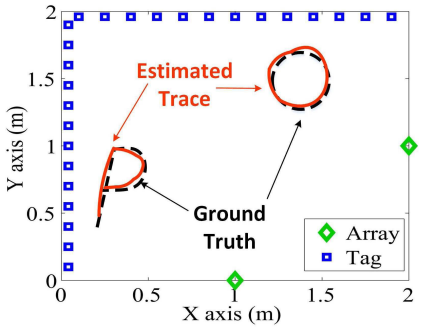


Fig. 22. Passively track the fist's writing in the air.

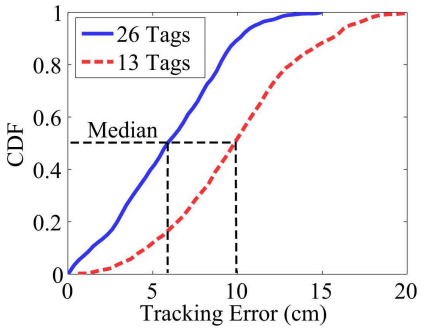


Fig. 23. Fist tracking accuracy of D-Watch.

VIII. Related Work

Device-Free Localization

Early device-free localization works rely on visible light camera [27], [29] and infrared sensor [18], [20]. Camera-based methods heavily reply on lighting conditions and have severe privacy issue. Infrared-based methods have difficulties to penetrate the walls. On the other hand, low frequency RF waves can penetrate walls easily [6], [7], [15], [57] and RF infrastructures are widely available. Thus, there are growing interests in exploring RF signals for device-free localization. Among RF-based methods, most device-free localization systems are RSSI/CSI fingerprint-based [9], [10], [41]–[43], [50]. They translate the localization problem to a fingerprint matching problem while each location is associated with a unique RSSI or CSI fingerprint [50]. The feasibility has been demonstrated for different technologies including RFID [17], [57], Wi-Fi [33], [50] and ZigBee [9], [60]. However, fingerprint-based methods need a large amount of human efforts to acquire and update the fingerprint database. Changes in the environment, such as the movements of furniture, will change the fingerprints [52], causing mismatches between the database and the new measurements. D-Watch on the other hand, does not need any labor-intensive training efforts, and only requires several baseline AoA measurements which can be carried out automatically without human interventions within seconds.

Recently, a lot of RSSI model-based device-free localization systems were proposed in order to reduce or even avoid the offline training efforts, such as RFID-based Tadar [57] and Twins [17] systems, Wi-Fi-based LiFS [44] and ACE [33] systems, ZigBee-based RTI [49] and RASS [60] systems, etc. The basic idea of these approaches is to model the wireless propagation channels mathematically and then estimate the target location from the distorted wireless signal. However, due to the complicated multipath situation in indoor environment, mathematical models do not fit well and the location accuracy is hence coarse. Also as the models focus on the LoS path, a dense transceiver deployment is usually needed to cover the area. Moreover, they require the prior knowledge of the locations of the transceivers, which are sometimes impractical. In contrast, D-Watch utilizes both the LoS path and the reflection paths for localization purpose, which significantly increases the coverage area and accuracy. The deployment density can thus be reduced, making it a promising candidate for large scale deployment. Further, D-Watch does not need to know the locations of the signal sources (such as the RFID tags or mobile devices) so the proposed system has a high flexibility for real life deployment.

Researchers also proposed some fine-grained device-free localization systems, such as Wi-Vi [7], Witrack [6], mtrack [48], etc. These systems though being able to achieve a high accuracy, require either dedicated signal (frequency modulated carrier wave, FMCW), specialized hardware (USRPs, WARP) or a very large bandwidth (60 GHz). Unlike these approaches, D-Watch is built on top of low cost COTS RFID devices and efficiently utilizes the "detrimental" multipaths to improve the coverage rate and localization accuracy.

Phase Calibration

Besides the wired phase calibration method introduced in ArrayTrack [52], some wireless phase calibration methods have also been proposed in Argos [36] and Phaser [16]. The method proposed in Argos can not be applied directly to our case as the RFID reader does not support transmission from one antenna on the reader to the other antennas. Also Argos requires all the transceivers to cooperate with each other for calibration, which interrupts the ongoing data transmissions. To overcome this limitation, Phaser is proposed and can carry out auto-calibration without interrupting the ongoing transmission. However, the calibration accuracy of Phaser is coarse. Our proposed calibration method not only achieves a much higher accuracy but also does not interrupt the ongoing data communication.

IX. Conclusion

D-Watch is the first device-free localization system that utilizes both the direct path and multipaths to provide decimeter-level localization accuracy without offline training. We propose a wireless phase calibration scheme to remove the random phase offsets at the radio front ends and a novel power MUSIC algorithm to accurately detect the angle information of the target. Comprehensive real-world experiments in different environments demonstrate the effectiveness of D-Watch.

References

[1] *Alien Tags*. Accessed: May 7, 2016. [Online]. Available: http://www.alientechnology.com/tags/

[2] *Ans-900 RFID Antenna*. Accessed: May 7, 2016. [Online]. Available: http://rf-links.com/newsite/pdf/ans-900.pdf

[3] *Epc Gen2, Epcglobal*. Accessed: May 7, 2016. [Online]. Available: www.gs1.org/epcglobal

[4] Impinj Inc. Accessed: May 7, 2016. [Online]. Available: http://www.impinj.com/products/readers/speedway-revolution/

[5] *Q900f-900 RFID Antenna*. Accessed: May 7, 2016. [Online]. Available: www.hrtantenna.com/en/products

[6] F. Adib, Z. Kabelac, and D. Katabi, "Multi-person localization via RF body reflections," in *Proc. Usenix NSDI*, 2015, pp. 279–292.

[7] F. Adib and D. Katabi, "See through walls with WiFi!" in *Proc. ACM SIGCOMM*, vol. 43. 2013, pp. 75–86.

[8] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*. Hoboken, NJ, USA: Wiley, 2006.

[9] L. Chang *et al.*, "FALE: Fine-grained device free localization that can adaptively work in different areas with little effort," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 45, no. 5, pp. 601–602, 2015.

[10] L. Chang *et al.*, "Fitloc: Fine-grained and low-cost device-free localization for multiple targets over various areas," *IEEE/ACM Trans. Netw.*, vol. 25, no. 4, pp. 1994–2007, Aug. 2017.

[11] B. Chen, V. Yenamandra, and K. Srinivasan, "Tracking keystrokes using wireless signals," in *Proc. ACM Mobisys*, 2015, pp. 31–44.

[12] Z. Zhou, B. Chen, and H. Yu, "Understanding RFID counting protocols," *IEEE/ACM Trans. Netw.*, vol. 24, no. 1, pp. 312–327, Feb. 2016.

[13] R. Diamant, H.-P. Tan, and L. Lampe, "LOS and NLOS classification for underwater acoustic localization," *IEEE Trans. Mobile Comput.*, vol. 13, no. 2, pp. 311–323, Feb. 2014.

[14] F. Thiesse, C. Floerkemeier, M. Harrison, F. Michahelles, and C. Roduner, "Technology, standards, and real-world deployments of the EPC network," *IEEE Internet Comput.*, vol. 13, no. 2, pp. 36–43, Mar./Apr. 2009.

[15] M. Flores, U. Klarman, and A. Kuzmanovic, "Wi-FM: Resolving neighborhood wireless network affairs by listening to music," in *Proc. IEEE ICNP*, Nov. 2005, pp. 43–53.

[16] J. Gjengset, J. Xiong, G. McPhillips, and K. Jamieson, "Phaser: Enabling phased array signal processing on commodity WiFi access points," in *Proc. ACM MobiCom*, 2014, pp. 153–164.

[17] J. Han *et al.*, "Twins: Device-free object tracking using passive tags," in *Proc. IEEE INFOCOM*, Apr. 2014, pp. 469–476.

[18] D. Hauschildt and N. Kirchhof, "Advances in thermal infrared localization: Challenges and solutions," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Sep. 2010, pp. 1–8.

[19] P. Jain, J. Manweiler, and R. R. Choudhury, "Overlay: Practical mobile augmented reality," in *Proc. ACM Mobisys*, 2015, pp. 331–344.

[20] J. Kemper and D. Hauschildt, "Passive infrared localization with a probability hypothesis density filter," in *Proc. IEEE Workshop Positioning Navigat. Commun. (WPNC)*, Mar. 2010, pp. 68–76.

[21] M. Khodja, K. Abed-Meraim, and A. Belouchrani, "Performance analysis for time-frequency MUSIC algorithm in presence of both additive noise and array calibration errors," *EURASIP J. Adv. Signal Process.*, vol. 1, no. 94, pp. 1–11, 2012.

[22] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, "SpotFi: Decimeter level localization using WiFi," in *Proc. ACM SIGCOMM*, 2015, pp. 269–282.

[23] T.-W. Kuo, K.-C. Lee, K. C.-J. Lin, and M.-J. Tsai, "Leader-contention-based user matching for 802.11 multiuser MIMO networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4389–4400, Aug. 2014.

[24] X. Li *et al.*, "Dynamic-MUSIC: Accurate device-free indoor localization," in *Proc. ACM UbiComp*, 2016, pp. 196–207.

[25] LIBSVM. *Library to Using SVM*. Accessed: May 7, 2016. [Online]. Available: www.csie.ntu.edu.tw/cjlin/libsvm/

[26] C. Luo *et al.*, "Pallas: Self-bootstrapping fine-grained passive indoor localization using WiFi monitors," *IEEE Trans. Mobile Comput.*, vol. 16, no. 2, pp. 466–481, Feb. 2017.

[27] H. Ma, C. Zeng, and C. X. Ling, "A reliable people counting system via multiple cameras," *ACM Trans. Intell. Syst. Technol.*, vol. 3, no. 2, pp. 67–83, 2012.

[28] T. McConaghy, E. Vladislavleva, and R. Riolo, "Genetic programming theory and practice 2010: An introduction," in *Proc. GECCO Companion Publication Annu. Genet. Evol. Comput. Conf.*, 2010, vol. 78. no. 1, pp. 3015–3056.

[29] R. Mohedano, A. Cavallaro, and N. García, "Camera localization using-trajectories and maps," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 4, pp. 684–697, Apr. 2014.

[30] F. M. Naini, J. Unnikrishnan, P. Thiran, and M. Vetterli, "Where you are is who you are: User identification by matching statistics," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 2, pp. 358–372, Feb. 2016.

[31] S. J. Orfanidis, *Electromagnetic Waves and Antennas*. Piscataway, NJ, USA: Rutgers Univ. Press, 2002.

[32] B. Partov, D. J. Leith, and R. Razavi, "Utility fair optimization of antenna tilt angles in LTE networks," *IEEE/ACM Trans. Netw.*, vol. 23, no. 1, pp. 175–185, Feb. 2015.

[33] I. Sabek, M. Youssef, and A. V. Vasilakos, "ACE: An accurate and efficient multi-entity device-free WLAN localization system," *IEEE Trans. Mobile Comput.*, vol. 14, no. 2, pp. 261–273, Feb. 2015.

[34] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 276–280, Mar. 1986.

[35] M. Shahzad and A. X. Liu, "Fast and accurate estimation of RFID tags," *IEEE/ACM Trans. Netw.*, vol. 23, no. 1, pp. 241–254, Feb. 2015.

[36] C. Shepard *et al.*, "Argos: Practical many-antenna base stations," in *Proc. ACM MobiCom*, 2012, pp. 53–64.

[37] Y. Shu, K. G. Shin, T. He, and J. Chen, "Last-mile navigation using smartphones," in *Proc. ACM MobiCom*, 2015, pp. 512–524.

[38] L. Sun, S. Sen, D. Koutsonikolas, and K.-H. Kim, "WiDraw: Enabling hands-free drawing in the air on commodity WiFi devices," in *Proc. ACM MobiCom*, 2015, pp. 77–89.

[39] J. Teng, B. Zhang, J. Zhu, X. Li, D. Xuan, and Y. F. Zheng, "EV-Loc: Integrating electronic and visual signals for accurate localization," *IEEE/ACM Trans. Netw.*, vol. 22, no. 4, pp. 1285–1296, Aug. 2014.

[40] D. Vasisht, S. Kumar, and D. Katabi, "Decimeter-level localization with a single WiFi access point," in *Proc. Usenix NSDI*, 2016, pp. 165–178.

[41] J. Wang *et al.*, "Transferring compressive-sensing-based device-free localization across target diversity," *IEEE Trans. Ind. Electron.*, vol. 62, no. 4, pp. 2397–2409, Apr. 2015.

[42] J. Wang *et al.*, "LCS: Compressive sensing based device-free localization for multiple targets in sensor networks," in *Proc. IEEE INFOCOM*, Apr. 2013, pp. 145–149.

[43] J. Wang *et al.*, "E-HIPA: An energy-efficient framework for high-precision multi-target-adaptive device-free localization," *IEEE Trans. Mobile Comput.*, vol. 16, no. 3, pp. 716–729, Mar. 2017.

[44] J. Wang *et al.*, "LiFS: Low human-effort, device-free localization with fine-grained subcarrier information," in *Proc. ACM MobiCom*, 2016, pp. 243–256.

[45] J. Wang and D. Katabi, "Dude, where's my card?: RFID positioning that works with multipath and non-line of sight," in *Proc. ACM SIGCOMM*, vol. 43. 2013, pp. 51–62.

[46] J. Wang, D. Vasisht, and D. Katabi, "RF-IDraw: Virtual touch screen in the air using RF signals," in *Proc. ACM SIGCOMM*, vol. 44. 2014, pp. 235–246.

[47] J. Wang *et al.*, "TagScan: Simultaneous target imaging and material identification with commodity RFID devices," in *Proc. ACM MobiCom*, 2017, pp. 1–14.

[48] T. Wei and X. Zhang, "mTrack: High-precision passive tracking using millimeter wave radios," in *Proc. ACM MobiCom*, 2015, pp. 117–129.

[49] J. Wilson and N. Patwari, "See-through walls: Motion tracking using variance-based radio tomography networks," *IEEE Trans. Mobile Comput.*, vol. 10, no. 5, pp. 612–621, May 2011.

[50] J. Xiao, K. Wu, Y. Yi, L. Wang, and L. M. Ni, "Pilot: Passive device-free indoor localization using channel state information," in *Proc. IEEE Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Jun. 2013, pp. 236–245.

[51] J. Xiong and K. Jamieson, "Towards fine-grained radio-based indoor location," in *Proc. ACM Workshop Mobile Comput. Syst. Appl.*, 2012, pp. 13–18.

[52] J. Xiong and K. Jamieson, "ArrayTrack: A fine-grained indoor location system," in *Proc. USENIX NSDI*, 2013, pp. 71–84.

[53] J. Xiong, K. Jamieson, and K. Sundaresan, "Synchronicity: Pushing the envelope of fine-grained localization with distributed MIMO," in *Proc. ACM Workshop Hot Topics Wireless*, 2014, pp. 43–48.

[54] J. Xiong, K. Sundaresan, and K. Jamieson, "ToneTrack: Leveraging frequency-agile radios for time-based indoor wireless localization," in *Proc. ACM MobiCom*, 2015, pp. 537–549.

[55] C. Xu *et al.*, "SCPL: Indoor device-free multi-subject counting and localization using radio signal strength," in *Proc. ACM/IEEE IPSN*, Apr. 2013, pp. 79–90.

[56] L. Yang, Y. Chen, X.-Y. Li, C. Xiao, M. Li, and Y. Liu, "Tagoram: Real-time tracking of mobile RFID tags to high precision using COTS devices," in *Proc. ACM MobiCom*, 2014, pp. 237–248.

[57] L. Yang, Q. Lin, X. Li, T. Liu, and Y. Liu, "See through walls with COTS RFID system," in *Proc. ACM MobiCom*, 2015, pp. 1–12.

[58] H.-C. Yen and C.-C. Wang, "Cross-device Wi-Fi map fusion with Gaussian processes," *IEEE Trans. Mobile Comput.*, vol. 16, no. 1, pp. 44–57, Jan. 2017.

[59] S. Yun, Y.-C. Chen, and L. Qiu, "Turning a mobile device into a mouse in the air," in *Proc. ACM Mobisys*, 2015, pp. 15–29.

[60] D. Zhang, Y. Liu, X. Guo, and L. M. Ni, "RASS: A real-time, accurate, and scalable system for tracking transceiver-free objects," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 5, pp. 996–1008, May 2013.

[61] Y. Zhu, Y. Zhu, B. Y. Zhao, and H. Zheng, "Reusing 60 GHz radios for mobile radar imaging," in *Proc. ACM MobiCom*, 2015, pp. 103–116.

Ju Wang received the Ph.D. degree in information engineering from Northwest University, Xi'an, China, in 2017.

His research interests include localization and signal processing.

Jie Xiong received the M.Sc. degree from the University College London and the Ph.D. degree from the Duke University. He is currently an Assistant Professor with the School of Information Systems, Singapore Management University. He was a recipient of the Google Doctoral Fellowship.

His research interests include wireless networking and mobile computing.

Hongbo Jiang received the B.S. and M.S. degrees from the Huazhong University of Science and Technology, China, and the Ph.D. degree from Case Western Reserve University in 2008. After that he joined the Faculty of Huazhong University of Science and Technology, where he is currently a Full Professor.

His research interests include computer networking, especially algorithms and architectures for wireless networks and mobile computing.

Xiaojiang Chen received the Ph.D. degree in computer software and theory from Northwest University, Xi'an, China, in 2010. He is currently a Professor with the School of Information Science and Technology, Northwest University.

His current research interests include localization and performance issues in wireless ad hoc, mesh, and sensor networks and named data networks.

Dingyi Fang received the B.S. and M.S. degrees in computer science from Northwest University, Xi'an, China, in 1983 and 1988, respectively, and the Ph.D. degree in computer application technology from Northwestern Polytechnical University, China, in 2001. He is currently a Professor with the School of Information Science and Technology, Northwest University.

His current research interests include mobile and distributed computing systems, network and information security, and wireless sensor networks.